

# Request Form for Translation

Translation Branch  
The world of foreign-prior art to you.



U. S. Serial No.: 09/749,743

148

Requester's Name: Tim Lao

Phone No.: 305-8955

Fax No.: \_\_\_\_\_

Office Location: CPK2 / 3 / C16

Art Unit/Org.: 2655

Group Director: \_\_\_\_\_

Is this for Board of Patent Appeals? \_\_\_\_\_

Date of Request: 10/15/03

Date Needed By: 10/29/03

(Please do not write ASAP-indicate a specific date)

PTO 2004-0273

S.T.I.C. Translations Branch

Phone: 308-0881  
Fax: 308-0989  
Location: Crystal Plaza 3/4  
Room 2C01

## SPE Signature Required for RUSH:

### Document Identification (Select One):

\*\* (Note: Please attach a complete, legible copy of the document to be translated to this form)\*\*

1. ☒ Patent Document No. 10-222332  
Language Japanese  
Country Code JP  
Publication Date Aug. 21, 1998  
No. of Pages \_\_\_\_\_ (filled by STIC)

2. ☐ Article Author \_\_\_\_\_  
Language \_\_\_\_\_  
Country \_\_\_\_\_

3. ☐ Other Type of Document \_\_\_\_\_  
Country \_\_\_\_\_  
Language \_\_\_\_\_

Document Delivery (Select Preference): E. mail  
☒ Delivery to Exmr. Office/Mailbox Date: 11-4-03 (STIC Only)

☐ Call for Pick-up Date: \_\_\_\_\_ (STIC Only)

To assist us in providing the most cost effective service, please answer these questions:

Will you accept an English Language Equivalent?  
Yes (Yes/No)

Will you accept an English abstract?  
No (Yes/No)

Would you like a consultation with a translator to review the document prior to having a complete written translation?  
No (Yes/No)

Check here if Machine Translation is not acceptable:  
(It is the default for Japanese Patents, '93 and onwards with avg. 5 day turnaround after receipt)  
☒

## STIC USE ONLY

### Copy/Search

Processor: 165  
Date assigned: 10/16/03  
Date filled: \_\_\_\_\_  
Equivalent found: see Attached (Yes/No)

Doc. No.: \_\_\_\_\_  
Country: \_\_\_\_\_

Remarks: \_\_\_\_\_

### Translation

Date logged in: 10-16-03  
PTO estimated words: 11,075  
Number of pages: 32  
In-House Translation Available: \_\_\_\_\_  
In-House: \_\_\_\_\_ Contractor: \_\_\_\_\_  
Translator: \_\_\_\_\_ Name: MC  
Assigned: \_\_\_\_\_ Priority: 10-20-03  
Returned: \_\_\_\_\_ Sent: 11-4-03  
Returned: 11-4-03

**BEST AVAILABLE COPY**

PTO 04-0273

Japanese Kokai Patent Application  
No. Hei 10[1998]-222337

COMPUTER SYSTEM

Yoshitaka Sawada et al.

UNITED STATES PATENT AND TRADEMARK OFFICE  
WASHINGTON, D.C. NOVEMBER 2003  
TRANSLATED BY THE RALPH MCELROY TRANSLATION COMPANY

JAPANESE PATENT OFFICE  
PATENT JOURNAL (A)  
KOKAI PATENT APPLICATION NO. HEI 10[1998]-222337

Int. Cl. <sup>6</sup> :	G 06 F 3/14 3/16 G 10 L 3/00
Filing No.:	Hei 9[1997]-28230
Filing Date:	February 13, 1997
Publication Date:	August 21, 1998
No. of Claims:	10 (Total of 12 pages; OL)
Examination Request:	Not filed

COMPUTER SYSTEM

[Konputa shisutemu]

Inventors:	Yoshitaka Sawada et al.
Applicant:	000006105 Meidensha Corporation

[There are no amendments to this patent.]

Claims

/2\*

1. A computer system, characterized by the fact that in a computer system equipped with a speech recognition interface that has a GUI environment, obtains a keyword by the speech recognition of an operation instruction through speech, and converts the keyword into a control command of a system software by decoding the keyword, the above-mentioned system software is equipped with a window switching means that switches the window corresponding to said control command to a window in an active state when the control command in which specific

---

\* [Numbers in right margin indicate pagination of the original text.]

decoded speech [i.e., words] are received from the above-mentioned speech recognition interface.

2. The computer system of Claim 1, characterized by the fact that the above-mentioned system software is equipped with a means that manages the window information of several applications, switches said window information to the newest active window of said application when each application is designated by the control command from the above-mentioned speech recognition interface, and switches it in order of newer windows of said applications by speaking the same application name.

3. The computer system of Claim 1, characterized by the fact that the above-mentioned system software is equipped with a means that displays the window activated by the control command from the speech recognition interface differently from other windows and displays a task list of each window.

4. The computer system of Claim 1 or 3, characterized by being equipped with a task switching means that sets the speech recognition interface to an input device in one of several applications by the control command from the above-mentioned speech recognition interface and sets a device other than said speech recognition interface to an input device in another application.

5. The computer system of Claim 1 or 2, characterized by the fact that the above-mentioned speech recognition interface is equipped with a means that prepares words for each application and a common word in a word dictionary for speech recognition and gives the above-mentioned system software a control command for displaying a keyword list registered in the above-mentioned word dictionary or the current effective keyword list on a screen when a specific keyword is given.

6. The computer system of any of Claims 1-5, characterized by the fact that the above-mentioned speech recognition interface has a table for converting the above-mentioned keyword into a control command; and the table makes several keywords correspond to one control command and constitutes table data for obtaining the control command by the reference to any of the keywords.

7. The computer system of any of Claims 1-5, characterized by the fact that the above-mentioned speech recognition interface is equipped with a means that displays keyword candidates being obtained by compensating the part after a prolonged sound from a speech input, in which the suffix of the known part of a keyword is extended, as a list.

8. The computer system of any of Claims 1-5, characterized by the fact that the above-mentioned speech recognition interface is equipped with a means that repeatedly generates the control command corresponding to a keyword in accordance with the length of a prolonged

sound when speech accompanied with said prolonged sound in which the suffix of said keyword is extended or speech in which the prolonged sound is added to the keyword is recognized.

9. The computer system of any of Claims 1-8, characterized by the fact that the above-mentioned speech recognition interface is equipped with a speech operation button that implements speech recognition only when a user inputs speech, an input level check section that implements the speech recognition only when an actual input speech is at a level suitable for recognition or higher, and a recognition result reliability check section that checks the reliability of the speech recognition result and only accepts the accurately recognized speech.

10. The computer system of Claim 1 or 3, characterized by the fact that the above-mentioned system software is equipped with a means that adopts the window of the position being designated by a pointing device as a highly noticeable window by a user and raises the frame rate of said window.

#### Detailed explanation of the invention

[0001]

##### Technical field of the invention

The present invention pertains to a computer system equipped with a man-machine interface that has a GUI environment and receives a data input or operation instruction through speech.

[0002]

##### Prior art

A man-machine interface for interacting with a computer system where a user inputs or instructs data from the user to the computer system by an input device such as keyboard or pointing device (mouse, etc.) and displays or replies various kinds of data from the computer system to the user by an output device such as display.

[0003]

In addition to the man-machine interface device, input and output devices using speech have recently been developed. In a computer system, a speech synthesizer and a speech generator are installed as speech output devices, and speech recognizer and decoder are installed as speech input devices.

[0004]

Problems to be solved by the invention

In case data are input or instructed to the computer system through speech, a user speaks a predetermined keyword (phrase or word), and the computer system recognizes the speech and decodes it.

[0005]

Here, in case the computer system has a graphical user interface (GUI) environment such as X-Window system or MS-Windows, several applications (windows) can be processed on the same screen, however in actuality, only one application (window) is active.

/3

[0006]

In such a computer system, in case the user operates a certain application through speech, said application (window) may be in an active state, however if another application (window) is active, the active application [currently] being used by the user can no longer be used.

[0007]

In this case, in the prior art, the window to be operated was activated by a pointing device such as mouse, and a verbal operation instruction was given. Since the operation of the mouse, etc., was interposed, the operation efficiency deteriorates.

[0008]

Also, if the application to be opened is misunderstood, an irrelevant application is likely to be erroneously opened.

[0009]

Also, in the computer system having the GUI environment, a number of windows are opened and active, however since a number of windows overlap each other, sometimes, the user cannot understand active window, the kind of registered application, the task list of the applications, keyword list structure, etc. In this case, the upper window is also moved and erased by mouse, etc., and while moving the lower window to the upper area, and the active window is sought, which is complicated.

[0010]

Also, in the instruction of various kinds of operations such as application switching, control contents switching, or screen operation, the keywords corresponding to them are spoken,

and speech recognition and decoding are required. However, in case the keywords are forgotten or only partially forgotten, the operation cannot be initiated by speech.

[0011]

The purpose of the present invention is to provide a computer system that is equipped with a man-machine interface, which has a GUI environment and instructs the input and operation of data using speech, and also increases the operation efficiency by a reliable and easy speech-activated selection and operation of a window and an application.

[0012]

Another purpose of the present invention is to provide a computer system that easily determine a correct keyword, even in case the keyword being stored by a user is ambiguous or forgotten.

[0013]

Means to solve the problems

In order to achieve the above-mentioned purposes, the present invention is characterized by the fact that in a computer system equipped with a speech recognition interface that has a GUI environment, obtains a keyword by the speech recognition of an operation instruction through speech, and converts the keyword into a control command for the system software by decoding the keyword, the above-mentioned system software is equipped with a window switching means that switches the window corresponding to said control command to an active window when the control command in which specific speech is decoded is received from the above-mentioned speech recognition interface.

[0014]

According to this constitution, since a window is switched by a specific keyword by speech-activation, a window switching operation using a mouse, etc., is not required. The window switching includes a method that sequentially activates windows in a priority order, a method that uses different words/phrases in the switching start and end of the windows, starts an automatic switching control at each fixed time through the switching start word/phrase by the system, and stops the switching through the switching end word/phrase, and a method that combines both. Also, a window switching method that gives a "marker" with a priority order to each window and activates an optional window by designating the "marker" through speech is included. Furthermore, a window switching method that manages the window information of several applications, switches a window to the newest active window when each application is

designated by speech, and switches it in a new window sequence of said applications by the same speech is also included.

[0015]

Also, this system is characterized by the fact that the above-mentioned system software is equipped with a means that displays the window activated by the control command from the speech recognition interface differently from other windows and displays a task list of each window.

[0016]

According to this constitution, an active window is discriminated from other windows and displayed, and a task list in which an active application is discriminated is displayed, so that a user easily recognizes the active window and application.

[0017]

Also, this system is characterized by being equipped with a task switching means that sets the speech recognition interface to an input device in one of several applications by the control command from the above-mentioned speech recognition interface and sets a device other than said speech recognition interface to an input device in another application.

[0018]

According to this constitution, using the task switching, an input into an application through speech and an input into other applications through other input devices such as mouse are set, so that several windows can be operated [simultaneously].

[0019]

Also, this system is characterized by the fact that the above-mentioned speech recognition interface is equipped with a means that prepares words for each application and a common word in a word dictionary for speech recognition and gives the above-mentioned system software a control command for displaying a keyword list registered in the above-mentioned word dictionary or the current effective keyword list on a screen when a specific keyword is given.

[0020]

According to this constitution, when a user speaks a specific keyword such as keyword list, a keyword list registered in a word dictionary or the current effective keyword list is displayed on the screen, and a keyword required for a user can be detected from the keyword list being displayed on the screen. The keyword screen display includes a display by applications, a display of a hierarchical structure, and a keyword input through speech by including the abbreviation of a keyword, and a keyword display using a macro function.

/4

[0021]

Also, this system is characterized by the fact that the above-mentioned speech recognition interface has a table for converting the above-mentioned keyword into a control command; and the table makes several keywords correspond to one control command and constitutes table data for obtaining the control command by the reference of any of the keywords.

[0022]

According to this constitution, even if a user does not accurately remember a keyword, when a keyword close or related to it is spoken, the intended control command of the user can be obtained by the speech recognition and the control command-keyword conversion.

[0023]

Also, this system is characterized by the fact that the above-mentioned speech recognition interface is equipped with a means that displays keyword candidates being obtained by compensating the part after a prolonged sound from a speech input, in which the suffix of the known part of a keyword is extended, as a list.

[0024]

According to this constitution, a user extends the suffix of the known part of a keyword, so that keyword candidates are displayed on the screen. A correct keyword is determined from the screen display, and an operation can be initiated by using the correct keyword.

[0025]

Also, this system is characterized by the fact that the above-mentioned speech recognition interface is equipped with a means that repeatedly generates the control command corresponding to a keyword in accordance with the length of a prolonged sound when the speech accompanying said prolonged sound in which the suffix of said keyword is extended or a speech in which the prolonged sound is added to the keyword is recognized.

[0026]

According to this constitution, instead of repeatedly speaking the same continuous control command, one keyword accompanied with a prolonged sound is spoken, so that one keyword is spoken to repeat the same control.

[0027]

Also, this system is characterized by the fact that the above-mentioned speech recognition interface is equipped with a speech operation button that implements speech recognition only when a user inputs speech, an input level check section that implements the speech recognition only when actual input speech is at a level suitable for recognition or higher, and a recognition result reliability check section that checks the reliability of the speech recognition result and only accepts the accurately recognized speech.

[0028]

According to this constitution, a wrong recognition due to a noise penetration or insufficient level speech is prevented, and the reliability of the recognition result is also checked, so that a wrong recognition is prevented.

[0029]

Also, this system is characterized by the fact that the above-mentioned system software is equipped with a means that adopts the window being designated by a pointing device as a window with a high degree of notice for a user and raises the frame rate of said window.

[0030]

According to this constitution, a window with a high degree of user interest can be easily designated without operating an active window.

[0031]

Embodiments of the invention

(First embodiment)

Figure 1 is a block diagram showing a speech recognition interface section of an embodiment of the present invention. A microphone 1 is a speech input means for giving an operating instruction to the computer system by a user.

[0032]

A speech recognition interface part 2 has speech recognition part 3, word dictionary 4 for a speech recognition, keyword control command conversion part 5, and keyword-control command table 6.

[0033]

The speech recognition part 3 collates an audio signal from the microphone 1 with the word dictionary 4 for a speech recognition, and the keyword with a matched recognition is output as a result. For example, in the case a user speaks "open," a character string of "open" is obtained as a keyword.

[0034]

The word dictionary 4 for a speech recognition, as shown in the figure, prepares words for each application and a common word for each application being implemented on the system software, and among them, the words and the common word for the application in a currently active state are effective.

[0035]

The keyword control command conversion part 5 obtains the corresponding control command, referring to keyword-control command table 6, using the keyword from the speech recognition part 3.

[0036]

In the table 6, control commands which are one-to-one correspondent to each keyword are described, and the control command corresponding to the keyword "open" are substituted by the operation of keyboard, mouse, etc., and described.

[0037]

A system software 8 of a computer 7 is constituted by a software having a GUI environment such as MS-Window and X-Window system and has a function of responding to the control command from the keyword control command conversion part 5 in addition to a function of responding to the operation instruction from the keyboard, etc.

[0038]

A display 9 has a graphic display function corresponding to the display control from the system software 8, and a display screen 9A has a display control function for displaying several applications (windows) on the same screen.

[0039]

Here, in this embodiment, virtual window priority orders are given to all the windows being displayed on the display screen 9A in the display control means of the system software 8, and the currently active window has the highest priority order.

[0040]

Also, the window positioning behind the active window has the second priority order, and the window behind it is the third window. In this manner, the priority order corresponding to the window position is set.

[0041]

Then, the system software 8 has a window switching function that sets a certain specific keyword (for example, "window") and switches the window positioning behind the currently active application (window), that is, the window with the second window priority order, to an active window.

/5

[0042]

During window switching by the window switching function, the window which had been active previously has the lowest window priority order.

[0043]

Therefore, according to this embodiment, when an intended application is in an inactive state, a user speaks the window switching keyword, so that the speech is recognized by the speech recognition part 3 and converted into the control command by the keyword control command conversion part 5. Thereby, the window on the screen 9A is switched by the system software 8, and in this switching, each time the window switching keyword is spoken, the window is switched to a higher priority order. As a result, a desired window can be in an active state.

[0044]

Thus, the user may set the window to be operated to an active state simply by speaking a specific keyword without operating a pointing device such as mouse, and an operation can be instructed to the window in an active state through speech, so that the operation efficiency can be raised and an incorrect operation can be prevented.

[0045]

Also, data may be input by either a method using keyboard, etc., or a method that recognizes and decodes the data by the speech recognition interface section 2.

[0046]

Also, in switching of windows, each time a specific keyword (control command) is received, windows have been switched in order of a high priority, however the window switching method can be appropriately changed through the system software.

[0047]

For example, when a window switching start keyword is received, an active window is automatically switched at a certain time interval (for example, each 1 sec), and when a window switching stop keyword is received, the window switching is stopped. Furthermore, a combined method including the method of this embodiment can be adopted. Furthermore, a method that initiates a window switching command and a window number through speech and sets the window with said window number to an active window by the system may also be adopted.

[0048]

Also, "markers" with priority orders are given to each window being displayed on a graphic display, and the "marker" is designated by a word/phrase, so that an optional window can be activated.

[0049]

For example, in the setup or cancellation of "marker," a user generates a certain specific keyword, for example, "setup of marker 1" or "cancellation of marker 1," so that "marker 1" is set up or canceled on the currently active window. Then, the window to be activated is designated by speaking "marker 1," etc.

[0050]

In this case, the "marker" being set up and canceled is displayed as an icon, and it is preferable to detect the "marker" being set up and canceled at a glance. Also, the "marker" is used through speech, however its setup or cancellation can be done by other input means such as a keyboard.

[0051]

In the window switching by the above "marker," compared with the switching by the keyword of editor, etc., a desired window can be activated by a word/phrase through switching by the "marker" at a window unit.

[0052]

Also, a window management section is installed in the system software 8, the window information consisting of a window name being displayed, an active window name, and an active window history is managed by the management section, and the window can also be switched to the newest active window of each application.

[0053]

For example, in a state in which several applications A-D are started and non-priority windows A1 and A2 in each application are displayed, if "application B" is spoken, the management section searches for the newest active window of the application B and activates the application B1, for instance.

[0054]

Furthermore, if "application B" is continuously spoken, the next application B2 is activated.

[0055]

(Second embodiment)

In the above-mentioned embodiment, a window is switched through speech, and the corresponding application is operated. However, in case a number of windows overlap each other, a user hesitates in deciding if the application is in an active state. Also, the user sometimes cannot detect the existence of non-priority windows.

[0056]

Accordingly, in this embodiment, as shown in Figure 2, while several applications are started, for example, when "task list" is spoken, the system software 8 displays the application list being currently in operation (each window name) as a task list from the control command corresponding to the keyword as shown in Figure 3.

[0057]

At that time, a knitted shape is displayed or special color and frame are displayed so that the currently active application (for example, application A) may be easily understood. Also, in case a non-priority window exists in each window, for example, in an application D, its non-priority window is also simultaneously displayed by a hierarchical structure. Also, the active window can be switched in a task list display state. /6

[0058]

According to this embodiment, a user is not required to remember the currently active window, and in conforming the active window, several window names can be decided at a look and can be easily recognized by the speech of a specific keyword such as "task list."

[0059]

(Third embodiment)

Figure 4 shows another embodiment of the present invention. The difference between said figure and Figure 1 is that a task switching part 10 is provided.

[0060]

A task conversion part 10 switches and controls an input device for the system software 8.

[0061]

In a normal case, one window (for example, application A) is operated, however in this state, if "task switching B" is given to the interface 2 by a speech, the control command is given to the task switching part 10.

[0062]

At that time, the task switching part 10 switches the input device so that the application B may be input through speech and the application A can be input only from the keyword or mouse.

[0063]

In order to help the user understand the switching state of the input device, the applications A and B of the window display screen are different from the other applications, and even the application examples are differently displayed, for example, by different colors.

[0064]

The return from the above-mentioned input device switching state is carried out with a return command to the task switching part 10 through the interface 2 by giving the "task switching B" using speech.

[0065]

According to this embodiment, several windows can be [simultaneously] operated.

[0066]

(Fourth embodiment)

Figure 5 shows another embodiment. In this embodiment, when a specific keyword is given to the keyword control command conversion part 5, the keyword list registered in the word dictionary can be displayed and controlled as a list.

[0067]

For example, in the constitution of Figure 5, when a user speaks a speech of "keyword list" through the microphone 1, the speech recognition part 3 gives the speech recognition result to the keyword control command conversion part 5. Said conversion part 5 reads out the keyword group registered in the word dictionary 4, uses it as a keyword list data, gives the control command of the keyword list to the system software 8, and displays the keyword list on the screen 9A by the display control of the system software 8.

[0068]

Figure 6 shows a screen display example of the keyword list, and the keywords of all applications and a common word list are displayed. Also, a knitted shape is displayed so that the currently effective keyword may be understood.

[0069]

Also, in the case of each application, for example, a word list is displayed with a hierarchical structure for each individual application as shown in application A.

[0070]

Also, the display of the keyword list may be the keyword display of the current application.

[0071]

According to this embodiment, since the currently effective keyword or all the keywords are displayed in a list type, a user may remember the keyword for calling the keyword list without remembering the keywords.

[0072]

Also, since the keyword list with a hierarchical structure is displayed, the recognition of the keyword for an application is easy. Also, the keywords are easily related between the applications.

[0073]

(Fifth embodiment)

This embodiment, similar to the above-mentioned embodiments, is the case where a specific keyword is given to the keyword control command conversion part 5, and the currently effective keyword group is displayed and controlled as a list group.

[0074]

For example, when a user speaks the phrase related to the "keyword list" through the microphone 1, the currently effective keyword groups are read out of the word dictionary and displayed on the screen 9A. At that time, each corresponding abbreviation (number or symbol) is simultaneously displayed in the keyword groups.

[0075]

Then, from the keyword groups being displayed, similarly to the above-mentioned embodiments, an operation is instructed by a speech input method of the keyword or a speech input of the abbreviation (number or symbol) being added to the list display of the keyword group.

[0076]

Figure 7 shows an operation example of the case where "Fukuoka" monitoring situation is determined out using a certain monitoring control system software.

[0077]

First, when a user speaks "keyword list," the keyword control command conversion part 5 reads the currently effective keyword groups out of the word dictionary 4 and displays the control command and the keyword groups as a menu screen A1 using the system software 8.

[0078]

Next, from the menu screen A1 being displayed, the user speaks "file" to know the keyword of the file. Thus, the effective keyword groups for the file are read out of the word dictionary 4 and displayed as a file screen A2. Next, though it is possible to select "execution" from the file screen A2 by speaking "execution," in the figure, the keyword group display of the execution screen A3 is obtained by speaking "1" as an abbreviation. Similarly, a monitoring control screen A4 is obtained from the execution screen A3 by speaking "2" as an abbreviation of the monitoring control, and "Fukuoka" monitoring situation can be obtained by speaking "4" as an abbreviation for designating "Fukuoka" of the screen.

17

[0079]

Here, the keyword "macro" being displayed on each screen A1-A4 gives a function of macro-processing speaking words and a speaking sequence in a speech and can be implemented by speaking fewer words in a high-frequency operation.

[0080]

Figure 8 shows the case where a monitoring situation is obtained using the macro, and a macro screen A5 is obtained by a speech input of "11" as an abbreviation of the macro of the menu screen A1. On the macro screen A5, "Fukuoka" as a geographical name and "monitoring" as a division of the monitoring control are displayed at the same time, and a monitoring situation display of "Fukuoka" is obtained by a speech [input] of "3" as an abbreviation.

[0081]

As mentioned above, according to this embodiment, without preparing keyword lists and without the user remembering keywords, a required keyword can be easily detected from the screen display by remembering one keyword for calling a keyword list.

[0082]

Also, according to this embodiment, the abbreviations of the keywords are received in addition to the keywords themselves, so that the number of words spoken by a user can be

reduced, the complexity can be minimized, and the speech recognition rate of the system can be improved.

[0083]

Also, according to this embodiment, the spoken words and a speaking sequence in a speech are macro-processed, so that the operation can be simplified and sped up and an intended operation instruction of the user can be realized.

[0084]

(Sixth embodiment)

Figure 9 shows another embodiment of the present invention and shows table data of a keyword-control command table.

[0085]

In the above-mentioned embodiments up to now, since the window switching and the operation instruction in response to speaking a specific keyword are carried out, it is necessary for a user to remember the keyword. For this reason, when the memory of the user is indistinct and a correct keyword cannot be accurately remembered, the window switching and the operation instruction are impossible.

[0086]

In order to solve this problem, preparing keyword lists is considered, however in dynamically switching applications in the GUI environment, preparing keyword lists for each application is complicated and impractical. Also, each time the application is changed, the list has to be rewritten.

[0087]

Accordingly, in this embodiment, in order to obtain an intended control command of a user who has an indistinct memory without preparing keyword lists, the keyword-control command table 6 as shown in the figure is prepared.

[0088]

In the example of Figure 9, the case where several (three in the figure) keywords "aku," "ake," and "open" are prepared for the control command description to open file open is shown.

[0089]

Therefore, when a user is opening a file using speech, if "ake" or "open" are spoken instead of a correct keyword "aku," it is converted into a correct control command by the conversion part 5 referring to the keyword control command table 6, so that the intended control is realized. Similarly, windows are also switched by preparing several keywords.

[0090]

According to this embodiment, it is not necessary for a user to remember a correct keyword, and without preparing keyword lists, an operation instruction can be obtained with a prescribed keyword by speaking related words that can be associated with the operation.

[0091]

Also, editing such as the addition or change of several keywords to the table 6 can be implemented by the input from a keyboard, however the editing may also be realized through speech.

[0092]

In the editing through speech, for example, said table can be set to an editable state by giving "editing of the table 6" as a control command to the system software 8 through speech, a keyword (for example, a speech of "registration") of the keyword editing is spoken, and the part having said keyword is retrieved from the table 6 by speaking a correct keyword. Then, the keyword of another name is registered in the table by speaking the keyword of said another name.

[0093]

(Seventh embodiment)

In the embodiments up to now, it is necessary to accurately remember at least one keyword, however since the memory of a user is indistinct, a correct keyword is forgotten, so that only a portion of the keyword is sometimes known.

[0094]

Accordingly, in this embodiment, a system for obtaining an intended keyword from the halfway known part of the keyword is provided.

[0095]

In this embodiment, a user inputs a partial keyword through speech, and the suffix of the known part in the speech input is extended and spoken, so that a list of correct keyword candidates obtained by compensating the part after the prolonged sound is displayed.

[0096]

For example, in the case a keyword for displaying the contents of a file is "display," if "hyo" is known in the keyword "display (hyoji)" registered in the table 6, the suffix is extended and spoken like "hyo-."

[0097]

After receiving the keyword of which the suffix is a prolonged sound, the keyword control command conversion part 5 searches for the keyword in the table 6, gives a screen display command of a list of correct keyword candidates (for example, display, standard, evaluation, etc.) starting with the recognized keyword "hyo" to the system software 8, and obtains the screen display.

[0098]

From the screen display, the user once again speaks the correct keyword "display," so that the operation instruction of the correct keyword "display" can be obtained.

[0099]

According to this embodiment, even if the user forgets the suffix of a correct keyword, if the initial portion is remembered, the correct keyword is found out of the keyword list displayed by the system, so that the correct keyword can be input through speech.

/8

[0100]

(Eighth embodiment)

In the embodiments up to this point, in the speech input with a known correct keyword, in the case a continuous operation such as "enlargement" and "reduction" of the screen and "movement" in all directions as a keyword is instructed, the user repeatedly speaks the keyword (for example, "enlargement," "enlargement," and "enlargement") to obtain an expected screen state, and this is a complicated process.

[0101]

Accordingly, in this embodiment, a system for simplifying a keyword input of the same continuous operation instruction is provided.

[0102]

In this embodiment, in speaking a prolonged sound in which the suffix of a keyword is extended, the same continuous control command can be repeated in accordance with the length of the prolonged sound, so that the repeated speaking of the same keyword is not required.

[0103]

For example, in speaking the keyword "kakudai" for enlarging a screen, "kakuda-i" with a prolonged sound is spoken. For this speech, the speech recognition part 3 recognizes an enlargement operation from the part of the speech "kakuda," referring to the word dictionary 4, gives "kakudai" as a keyword and a code meaning its continuity to the keyword control command conversion part 5, and repeats the generation of the control command of the screen enlargement by said conversion part 5.

[0104]

Then, when "i" being the suffix of the keyword "kakudai" is recognized, the speech recognition part 3 gives a code meaning to stop of the repetition control to the conversion part 3, so that the conversion part 5 stops the generation of the control command of "kakudai." Thereby, the system software 8 stops the enlargement display control of the screen.

[0105]

The reduction of the screen and the movement in all directions are similarly carried out. Furthermore, this process can also be utilized in menu selection, etc., using a pull down menu. In other words, when an item in the pull down menu is selected, each keyword "ue" and "shita" of the vertical movement of a cursor are generated with a prolonged sound, so that the cursor is moved in the vertical direction as long as the prolonged sound continues. When the speech is finished, the cursor movement is stopped, and the item at the stop position can be selected.

[0106]

According to this embodiment, in a continuous operation instruction by the same keyword, the user speaks the keyword with a prolonged sound, so that a continuous control of screen enlargement, etc., is carried out as long as the prolonged sound continues and the speech is finished when a desired screen state is set. Thereby, the continuous control is stopped, and the

user can obtain an expected screen state. Thus, the user may speak the keyword once, and a continuous operation instruction is made easy.

[0107]

Also, in this embodiment, instead of the keyword with a prolonged sound, a keyword in which a prolonged sound is added to the suffix (for example, "kakuda-i") is spoken, and the operation instruction of said keyword is repeated as long as the prolonged sound continues, so that an equivalent operation effect is exerted.

[0108]

(Ninth embodiment)

In the system constitution of Figures 1, 4, and 5, generally, since 100% recognition rate of the speech recognition cannot be achieved, some erroneous operation of the system is expected.

[0109]

Accordingly, in this embodiment, a system for reliably preventing an erroneous operation, even if an error occurs during the speech recognition, is provided.

[0110]

Figure 10 shows a speech recognition interface of this embodiment. A speech operation button 11 is a software button, and an enabling signal for speech input is given to a speech input interface 12 by its ON operation. With the installation of the button 11, when the button 11 is turned off, the speech input is cut off, and the penetration of noise due to surrounding noise into the speech recognition part 3 is eliminated, so that an erroneous operation due to the noise is prevented.

[0111]

An input level check section 13 checks the level of the speech signal passed through the interface 12, and the speech signal with an input level lower than a normal level is cut off, so that an erroneous speaking due to the low-level speech recognition is prevented.

[0112]

Also, when the speech signal is cut off, the input level check section 13 transfers a signal indicating the cut-off to a response function part 14.

[0113]

A recognition result reliability check section 15 checks the reliability of the recognition result of the speech recognition 3 and stops the transfer of the recognition result to the conversion part 5 when the reliability is low. Thus, operations caused by erroneous speech are prevented by checking the reliability of the speech recognition.

[0114]

Also, when the reliability is low, the check section 15 transfers a signal indicating an operation stop to the response function part 14.

[0115]

When the speech recognition abnormality signal from the check section 13 or 15 is received, the response function part 14 requests a user to repeat the speech. Also, when the abnormality signal is received even after the request to repeat the speech, the speech operation button 11 is forcibly turned off, and the input request is made by the user or other input interfaces (mouse, keyboard, etc.)

[0116]

Therefore, according to this embodiment, only when the user inputs the speech by the speech operation button 11 is speech recognition is implemented, and as a result of check of the input level check section 13, only when the actual input speech is a level suitable for recognition or higher, is the speech recognition implemented. The reliability of the speech recognition result is checked by the recognition result reliability check section 15, and only the speech accurately recognized is responded to.

[0117]

Also, for the abnormality generation of the speech input and recognition, the user is urged to re-input the speech or to input the speech by another interface, so that reliable speech input and recognition can be obtained.

[0018]

(Tenth embodiment)

In simultaneously displaying several windows, a limited system resource can be optimally distributed by changing the frame rate of each window in accordance with the height of the degree of window (the degree of interest) of the user.

[0119]

The degree of interest can be estimated from the display state of the window. As its rule, as shown in Figure 11(a), it can be said that the degree of interest is high in the window closer to an active window and the smaller the area of the window, the lower the degree of interest. The frame rate of the window is also be changed by the degree of interest.

[0120]

In this rule, the degree of interest is automatically determined by the display state of the window, and in order for a user to optionally change the degree of interest, an operation such as a change of an active window is required, and the operation is complicated.

[0121]

Accordingly, in this embodiment, a system that can display an optimum distribution by reliably and automatically estimating the degree of interest of a user is provided.

[0122]

In this embodiment, as shown in Figure 11(b), the window in which a pointing device such as mouse is positioned is recognized as a window with a high degree of interest of a user by the system, and the frame rate of the window is raised. At that time, the active window is not changed.

[0123]

Therefore, in this embodiment, with the change of the pointer position of the pointing device such as mouse, the degree of interest of a window is automatically changed, and the frame rate is also changed. Thus, without a change operation of an active window, the user can optionally raise the degree of interest of a specific window. For this reason, the operation for raising the frame rate of the window with a degree of interest of the user is simplified, so that a more easily usable system can be constructed.

[0124]

Effect of the invention

As mentioned above, according to the present invention, since windows can be switched through speech using the speech recognition interface, a conventional window switching operation using mouse, etc., is not required, and the windows are switched through speech. Thus, the operation efficiency is raised, and erroneous operations can be prevented.

[0125]

Also, in the present invention, the window information of several applications is managed, and when each application is designated through speech, a window is switched to the newest active window of the application from the window information. The windows are switched in order of a new window of said application by the speech of the same application name. Thus, the window being often used is easily designated from the windows of the applications.

[0126]

Also, in the present invention, since an active window is discriminated from other windows and displayed by a speech input and a task list in which an active application is discriminated is displayed, a user easily recognizes the active window and application.

[0127]

Also, in the present invention, using a task switching, the input into an application by a speech and the input into other applications by other input devices such as mouse are carried out, so that several windows can be operated [simultaneously].

[0128]

Also, in the present invention, when a user speaks a specific keyword such as keyword lists, since the keyword registered in the word dictionary or the currently effective keyword list are displayed on a screen, the keyword required by the user can be easily detected from the screen display, so that all the effective keywords can be detected from the screen display, simply by remembering one keyword such as keyword lists.

[0129]

Also, in the present invention, since table data are constituted so that several keywords are made correspondent to one control command in the table for converting keywords into control commands and the control command is obtained from one keyword, even if a user cannot accurately remember the keyword, an intended control command of a user is obtained by a speaking a keyword related to it. Thus, an accurate memory of the keyword is not required, so that the usage method is improved.

[0130]

Also, in the present invention, a user extends the suffix of the known part of a keyword and speaks it, so that keyword candidates are displayed on a screen. Thus, the user can detect a

correct keyword from the screen display of the keyword candidates, and if part of the keyword is remembered, a correct keyword can be detected.

[0131]

Also, in the present invention, instead of repeatedly speaking the same continuous control command, one key word with a prolonged sound is spoken, so that the same control is repeated as long as the prolonged sound continues. Thus, an operation instruction by the same keyword is realized by one word in which the prolonged sound is added to said keyword, and the same continuous operation instruction is simplified.

[0132]

Also, in the present invention, only when a user carries out a speech input is the speech recognition implemented, and only when an actual input speech is a level suitable for the recognition or higher is the speech recognition implemented. The reliability of the speech recognition result is checked, and only the speech accurately recognized is responded to. Also, for the abnormality generation of the speech input and recognition, the user is urged to re-input the speech or to input the speech by another interface, so that reliable speech input and recognition can be obtained.

[0133]

Also, in the present invention, the window in which a pointing device such as mouse is positioned is recognized as a window with a high degree of interest by the system software, and the frame rate of the window is raised. Thus, without a change operation of an active window, the user can optionally raise the degree of interest of a specific window. Thereby, the operation /10  
for raising the frame rate of the window with a high degree of interest of the user is simplified, so that a more easily usable system can be constructed.

#### Brief description of the figures

Figure 1 is a block diagram showing an embodiment of the present invention.

Figure 2 is a window image in another embodiment.

Figure 3 is a task list window image in another embodiment.

Figure 4 is a block diagram showing another embodiment.

Figure 5 is a block diagram showing another embodiment.

Figure 6 is a keyword list window image in another embodiment.

Figure 7 shows an operation instruction sequence through speech in another embodiment.

Figure 8 shows an operation instruction sequence through speech in another embodiment.

Figure 9 shows a data constitution example of a keyword-control command table in another embodiment.

Figure 10 shows a constitutional example of a speech recognition interface in another embodiment.

Figure 11 is a window display example in accordance with the degree of interest in another embodiment.

Explanation of reference symbols:

- 1 Microphone
- 2 Speech recognition interface
- 3 Speech recognition part
- 4 Word dictionary
- 5 Keyword control command conversion part
- 6 Keyword-control command conversion part
- 7 Computer
- 8 System software
- 9 Display section
- 10 Task switching part
- 11 Speech operation button
- 12 Speech input interface
- 13 Input level check section
- 14 Response function part
- 15 Recognition result reliability check section

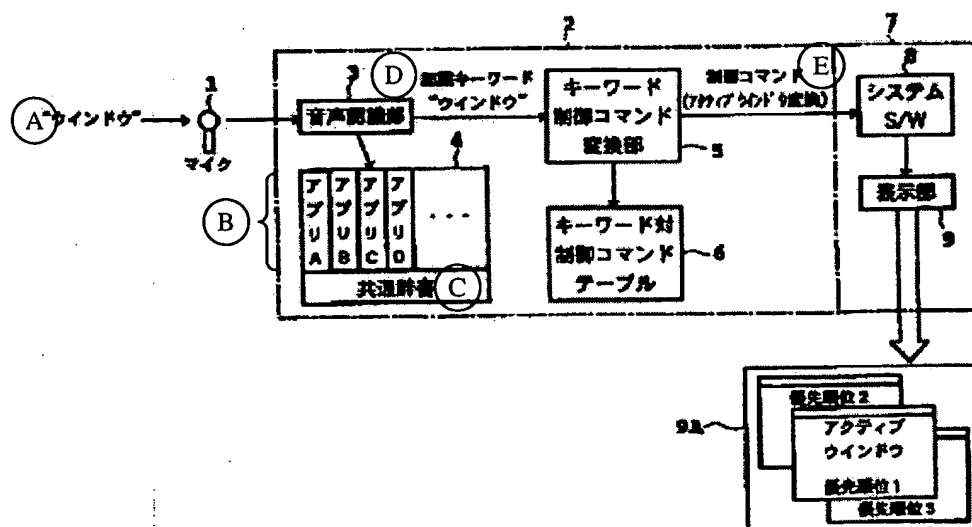


Figure 1. A block diagram showing an embodiment

- Key: A Window  
 B Application \_\_\_\_  
 C Common dictionary  
 D Recognition keyword "window"  
 E Control command (active window conversion)  
 1 Microphone  
 3 Speech recognition part  
 5 Keyword control command conversion part  
 6 Keyword-control command conversion part  
 8 System S/W  
 9 Display section  
 9A Priority order 2  
 Active window  
 Priority order 1  
 Priority order 3

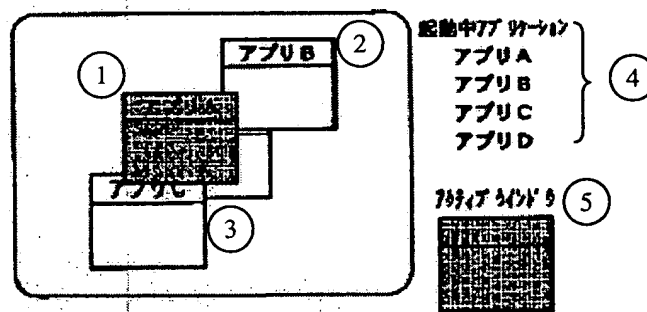


Figure 2. Window image

- Key: 1 Application A  
 2 Application B  
 3 Application C  
 4 Application being started  
 Application A  
 Application B  
 Application C  
 Application D  
 5 Active window

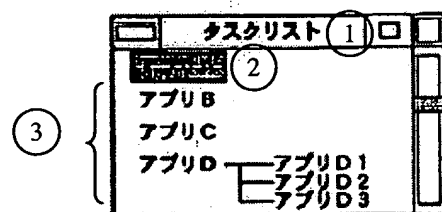


Figure 3. Task list window image

Key: 1 Task list  
 2 Application A  
 3 Application \_\_\_\_

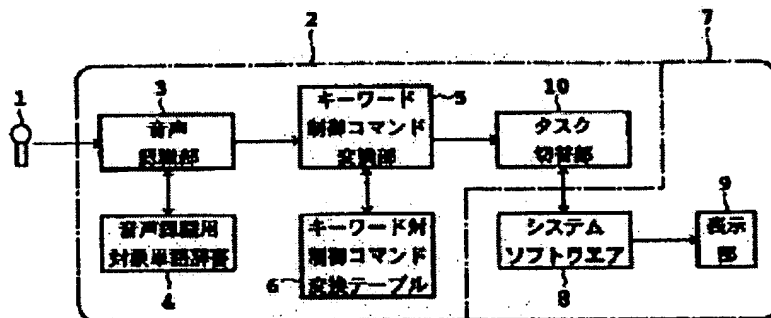


Figure 4. A block diagram showing another embodiment

Key: 3 Speech recognition part  
 4 Word dictionary  
 5 Keyword control command conversion part  
 6 Keyword-control command conversion part  
 8 System software  
 9 Display section  
 10 Task switching part

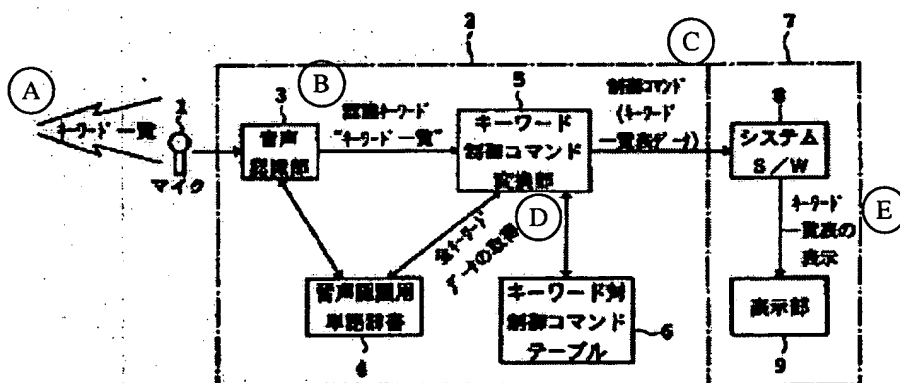


Figure 5. A block diagram showing another embodiment

Key: A Keyword list  
 B Recognition keyword  
 "Keyword list"  
 C Control command (keyword list data)  
 D Acquisition of all keyword data  
 E Display of keyword list  
 1 Microphone  
 3 Speech recognition part  
 4 Word dictionary for a speech recognition

- 5 Keyword control command conversion part  
 6 Keyword-control command conversion part  
 8 System S/W  
 9 Display section

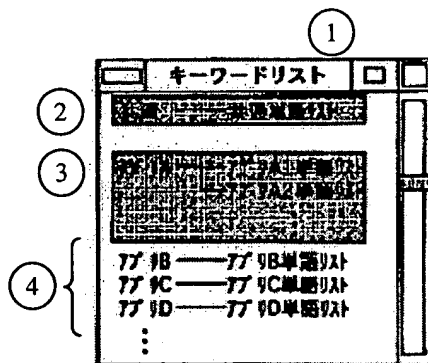


Figure 6. Keyword list window image

- Key: 1 Keyword list  
 2 Common Common word list  
 3 Application A Application A1 word list  
 Application A2 word list  
 4 Application B Application B word list  
 Application C Application C word list  
 Application D Application D word list

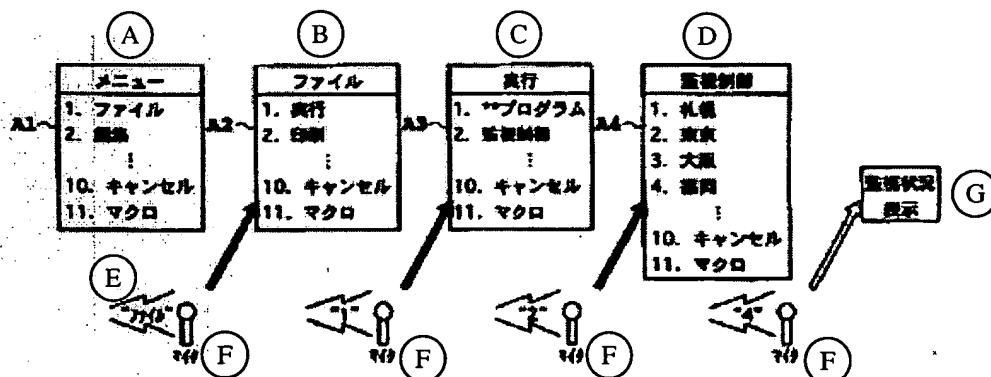


Figure 7. Operation instruction sequence through speech

- Key: A Menu  
 1 File  
 2 Edit  
 10 Cancel  
 11 Macro  
 B File  
 1 Execute  
 2 Print  
 10 Cancel  
 11 Macro  
 C Execute

- 1      \*\*Program  
 2      Monitoring control  
 10     Cancel  
 11     Macro  
 D      Monitoring control  
       1      Sapporo  
       2      Tokyo  
       3      Osaka  
       4      Fukuoka  
       10     Cancel  
       11     Macro  
 E      "File"  
 F      Microphone  
 G      Monitoring situation display

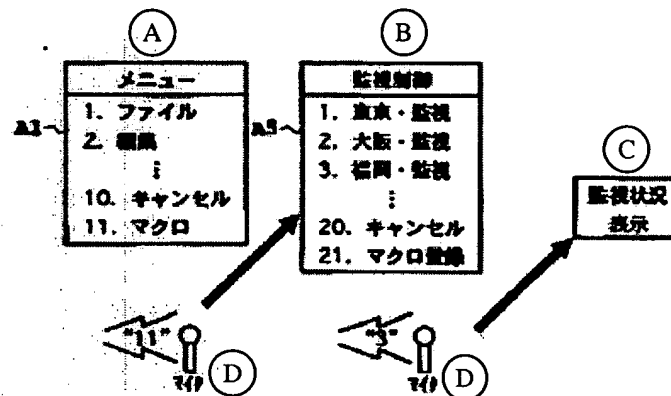


Figure 8. Operation instruction sequence through speech

- Key: A      Menu  
       1      File  
       2      Edit  
       10     Cancel  
       11     Macro  
 B      Monitoring control  
       1      Tokyo, monitoring  
       2      Osaka, monitoring  
       3      Fukuoka, monitoring  
       20     Cancel  
       21     Macro registration  
 C      Monitoring situation display  
 D      Microphone

①			②
キーワード1	キーワード2	キーワード3	制御コマンド記述
③ 開く	④ 開け	⑤ オープン	⑥ ファイルオープンのためのコマンドの記述
...	...	...	.....
...	...	...	.....
...	...	...	.....
...	...	...	.....

Figure 9. Table data constitution example

- Key: 1 Keyword \_\_\_\_  
 2 Control command description  
 3 Aku  
 4 Ake  
 5 Open  
 6 Command description for opening a file

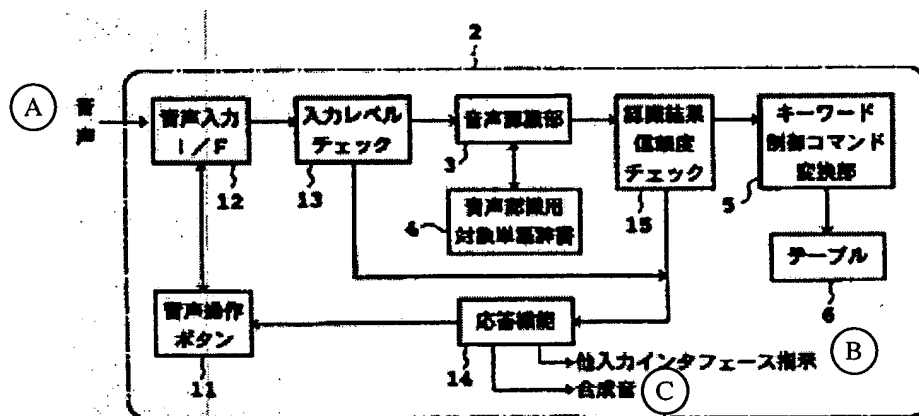


Figure 10. Constitutional example of a speech recognition interface system

- Key: A Speech  
 B Other input interface instruction  
 C Synthetic word  
 3 Speech recognition part  
 4 Word dictionary for voice recognition  
 5 Keyword control command conversion part  
 6 Table  
 11 Speech operation button  
 12 Speech input I/F  
 13 Input level check  
 14 Response function  
 15 Recognition result reliability check

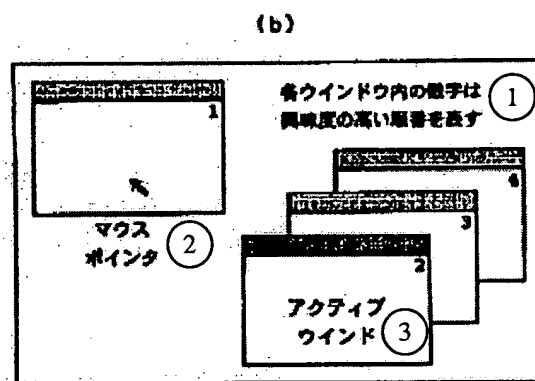
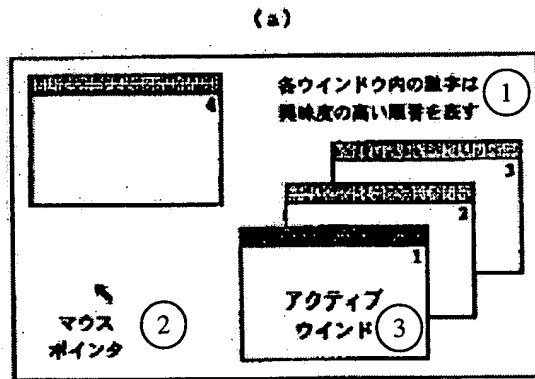


Figure 11. Decision example of the degree of interest

- Key:
- 1 The numerals in each window show the degree of interest in higher order.
  - 2 Mouse pointer
  - 3 Active window

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-222337

(43) 公開日 平成10年(1998) 8月21日

(51) Int.Cl. <sup>6</sup>	識別記号	FI
G 0 6 F 3/14	3 5 0	G 0 6 F 3/14 3 5 0 A
	3/16	3 2 0
G 1 0 L 3/00	5 5 1	G 1 0 L 3/00 5 5 1 Z

審査請求 未請求 請求項の数10 OL (全 12 頁)

(21) 出願番号 特願平9-28230

(22) 出願日 平成9年(1997) 2月13日

(71) 出願人 000006105

株式会社明電舎

東京都品川区大崎2丁目1番17号

(72) 発明者 沢田 喜正

東京都品川区大崎2丁目1番17号 株式会社明電舎内

(72) 発明者 佐々 学

東京都品川区大崎2丁目1番17号 株式会社明電舎内

(72) 発明者 木下 実

東京都品川区大崎2丁目1番17号 株式会社明電舎内

(74) 代理人 弁理士 志賀 富士弥 (外1名)

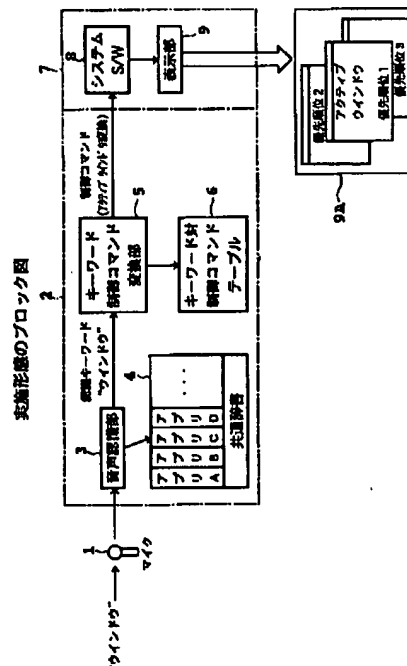
最終頁に続く

(54) 【発明の名称】 コンピュータシステム

(57) 【要約】

【課題】 GUI環境を持ち、音声によるデータ入力・操作指示を行うコンピュータシステムにおいて、音声によるウインドウ及びアプリケーションの確実・容易な選択・操作になり、操作効率も高める。

【解決手段】 マイク1からの音声を音声認識部3でキーワードとして認識し、このキーワードを変換部5でシステムソフトウェア8の制御コマンドに変換する音声認識インタフェース2とし、音声によるウインドウ切換え、アクティブなウインドウを他のウインドウから区別した表示、アクティブなアプリケーションを区別したタスクリスト表示、音声によるアプリケーションへの入力と、マウス等の他の入力デバイスによる他のアプリケーションへの入力とするタスク切換え、単語辞書4に登録されるキーワード又は現在有効なキーワードリストの画面表示を行う。キーワードの入力を簡易にする手段も含む。



## 【特許請求の範囲】

【請求項1】 GUI環境を持ち、音声による操作指示の音声認識によってキーワードを得、このキーワードを解読してシステムソフトウェアの制御コマンドに変換する音声認識インタフェースを備えたコンピュータシステムにおいて、前記システムソフトウェアは特定の音声を読み解した制御コマンドを前記音声認識インタフェースから受けたときに該制御コマンドに対応するウィンドウをアクティブな状態にするウィンドウ切替手段を備えたことを特徴とするコンピュータシステム。

【請求項2】 前記システムソフトウェアは、複数のアプリケーションのウィンドウ情報を管理し、前記音声認識インタフェースからの制御コマンドにより各アプリケーションを指定したときに該ウィンドウ情報から該アプリケーションの最新のアクティブウィンドウに切替え、同じアプリケーション名の音声により当該アプリケーションの新しいウィンドウ順に切替える手段を備えたことを特徴とする請求項1記載のコンピュータシステム。

【請求項3】 前記システムソフトウェアは、音声認識インタフェースからの制御コマンドによりアクティブにしたウィンドウを他のウィンドウとは異なるウィンドウの表示かつ各ウィンドウのタスクリスト表示を行う手段を備えたことを特徴とする請求項1記載のコンピュータシステム。

【請求項4】 前記音声認識インタフェースからの制御コマンドにより複数のアプリケーションのうち1つは音声認識インタフェースを入力デバイスとし、他の1つは該音声認識インタフェース以外のデバイスを入力デバイスとするタスク切替手段を備えたことを特徴とする請求項1又は3記載のコンピュータシステム。

【請求項5】 前記音声認識インタフェースは、音声認識のための単語辞書にはアプリケーション別の単語及び共通の単語を用意し、特定のキーワードが与えられたときに、前記単語辞書に登録されているキーワードリストを一覧表として画面表示、又は現在有効なキーワードリストを画面表示する制御コマンドを前記システムソフトウェアに与える手段を備えたことを特徴とする請求項1又は2記載のコンピュータシステム。

【請求項6】 前記音声認識インタフェースは、前記キーワードを制御コマンドに変換するテーブルを有し、このテーブルは1つの制御コマンドに複数のキーワードを対応付け、何れか1つのキーワードの参照で制御コマンドを得るテーブルデータ構成を特徴とする請求項1乃至5記載のうちの何れか1つに記載のコンピュータシステム。

【請求項7】 前記音声認識インタフェースは、キーワードの分かっている部分の語尾を延ばした発声入力から長音以降を補完して得られるキーワード候補をリスト表示する手段を備えたことを特徴とする請求項1乃至5記載のうちの何れか1つに記載のコンピュータシステム。

【請求項8】 前記音声認識インタフェースは、キーワードの語尾を延ばした長音を伴った音声、又はキーワードに長音を付加した音声を認識したときに、該長音の長さに応じて該キーワードに応じた制御コマンドを繰り返して発生する手段を備えたことを特徴とする請求項1乃至5記載のうちの何れか1つに記載のコンピュータシステム。

【請求項9】 前記音声認識インタフェースは、利用者が音声入力するときのみ音声認識を実行させる音声操作ボタンと、実際の入力音声を読み解に適当なレベル以上にあるときのみ音声認識を実行させる入力レベルチェック部と、音声認識結果の信頼度をチェックして正確に認識した音声のみを取り出す認識結果信頼度チェック部とを備えたことを特徴とする請求項1乃至8記載のうちのいずれか1つに記載のコンピュータシステム。

【請求項10】 前記システムソフトウェアは、ポインティングデバイスが指定する位置のウィンドウを利用者が注目する度合いの高いウィンドウとし、当該ウィンドウのフレームレートを高める手段を備えたことを特徴とする請求項1又は3記載のコンピュータシステム。

## 【発明の詳細な説明】

## 【0001】

【発明が属する技術分野】本発明は、GUI環境を持ち、音声によるデータ入力や操作指示を受け付けるマン・マシン・インタフェースを備えたコンピュータシステムに関する。

## 【0002】

【従来の技術】コンピュータシステムと利用者が対話を行うためのマン・マシン・インタフェースは、キーボードやポインティングデバイス（マウスなど）の入力デバイスにより人からコンピュータシステムへのデータの入力や指示を行い、ディスプレイ等の出力デバイスによりコンピュータシステムから人への各種データ表示や返答を行うようにしている。

【0003】これら、マン・マシン・インタフェース・デバイスによる他に、最近では音声により入出力を行うものがあり、コンピュータシステムには音声出力デバイスとして音声合成装置と音声発生装置を設け、音声入力デバイスとして音声認識装置と解読装置を設ける。

## 【0004】

【発明が解決しようとする課題】音声によってコンピュータシステムにデータの入力や指示を行う場合、あらかじめ決められたキーワード（言葉又は単語）を利用者が発声し、この音声をコンピュータシステムが音声認識し、その解読をする。

【0005】ここで、コンピュータシステムがX-Wind ow systemやMS-Windows等のグラフィカル・ユーザ・インタフェース（GUI）環境を持つ場合、同一画面上で複数のアプリケーション（ウィンドウ）を扱うことができるが、実際にアクティブなの

はそのうちの1アプリケーション(ウィンドウ)である。

【0006】この種のコンピュータシステムに対して、利用者があるアプリケーションに対して音声によって操作を行う場合、当該アプリケーション(ウィンドウ)がアクティブな状態であればよいが、他のアプリケーション(ウィンドウ)がアクティブになっていると利用者が操作使用とするアプリケーションに対して操作することができない。

【0007】このような場合、従来ではマウス等のポインティングデバイスによって操作したいウィンドウをアクティブにし、この後に音声による操作指示を行うようにしており、マウス等の操作が介在するため操作効率を悪くするものであった。

【0008】また、操作しようとするアプリケーションを間違えると、関係のないアプリケーションに対して誤操作してしまう恐れがある。

【0009】また、GUI環境を持つコンピュータシステムでは、多数のウィンドウを開いて操作を行うが、多数のウィンドウが重なり合うため、利用者はアクティブなウィンドウ及び登録アプリケーションの種別、アプリケーションのタスクリストやキーワードリスト構造などが分からなくなる場合がある。この場合もマウス等によって最上部のウィンドウを移動消去し、下層部のウィンドウを上層部に移動させながら操作対象ウィンドウを探すという煩わしさがある。

【0010】また、アプリケーション切替や制御内容の切替、また画面操作など、各種の操作指示にはそれに応じたキーワードを発声し、その音声認識と解釈を必要とするが、キーワードを忘れた場合や曖昧な記憶には音声による操作指示ができなくなる。

【0011】本発明の目的は、GUI環境を持ち、音声によるデータ入力・操作指示を行うマン・マシン・インタフェースを備えたコンピュータシステムにおいて、音声によるウィンドウ及びアプリケーションの確実・容易な選択・操作になり、操作効率も高めるコンピュータシステムを提供することにある。

【0012】本発明の他の目的は、利用者が記憶するキーワードが曖昧な場合や忘れた場合にも正しいキーワードの引き出しを容易にするコンピュータシステムを提供することにある。

【0013】

【課題を解決するための手段】本発明は、前記課題の解決を図るため、GUI環境を持ち、音声による操作指示の音声認識によってキーワードを得、このキーワードを解釈してシステムソフトウェアの制御コマンドに変換する音声認識インタフェースを備えたコンピュータシステムにおいて、前記システムソフトウェアは特定の音声を読み解した制御コマンドを前記音声認識インタフェースから受けたときに該制御コマンドに対応するウィンドウを

アクティブな状態にするウィンドウ切替手段を備えたことを特徴とする。

【0014】この構成により、音声による特定のキーワードによってウィンドウ切替を行うことにより、マウス等によるウィンドウ切替操作を不要にする。ウィンドウの切替には、優先順位をつけたウィンドウを順次アクティブにする方式、ウィンドウの切替開始と終了で異なる音声とすることで切替開始音声でシステム側が一定時間毎に自動的に切替える制御を開始し、切替終了音声で切替えを停止する方式、さらには両方式を組み合わせた方式を含む。また、各ウィンドウに優先順位を持たせた「しおり」を付け、この「しおり」を音声で指定することにより任意のウィンドウをアクティブにするウィンドウ切替方式も含む。さらに、複数のアプリケーションのウィンドウ情報を管理し、音声により各アプリケーションを指定するとその最新のアクティブウィンドウに切替え、さらに同じ音声で当該アプリケーションのうちの新しいウィンドウ順に切替えるウィンドウ切替方式も含む。

【0015】また、前記システムソフトウェアは、音声認識インタフェースからの制御コマンドによりアクティブにしたウィンドウを他のウィンドウとは異なるウィンドウの表示かつ各ウィンドウのタスクリスト表示を行う手段を備えたことを特徴とする。

【0016】この構成により、音声入力により、アクティブなウィンドウを他のウィンドウから区別した表示を行うこと及びアクティブなアプリケーションを区別したタスクリスト表示を行い、利用者がアクティブなウィンドウ及びアプリケーションを認識するのを容易にする。

【0017】また、前記音声認識インタフェースからの制御コマンドにより複数のアプリケーションのうち1つは音声認識インタフェースを入力デバイスとし、他の1つは該音声認識インタフェース以外のデバイスを入力デバイスとするタスク切替手段を備えたことを特徴とする。

【0018】この構成により、タスク切替により、音声によるアプリケーションへの入力と、マウス等の他の入力デバイスによる他のアプリケーションへの入力とし、複数のウィンドウを操作対象とすることができるようにする。

【0019】また、前記音声認識インタフェースは、音声認識のための単語辞書にはアプリケーション別の単語及び共通の単語を用意し、特定のキーワードが与えられたときに、前記単語辞書に登録されているキーワードリストを一覧表として画面表示、又は現在有効なキーワードリストを画面表示する制御コマンドを前記システムソフトウェアに与える手段を備えたことを特徴とする。

【0020】この構成により、利用者がキーワード一覧等の特定のキーワードを発声したときに単語辞書に登録されるキーワード又は現在有効なキーワードリストを画

5

面表示し、この画面表示されるキーワードリストから利用者が必要とするキーワードを知ることができるようにする。キーワードの画面表示にはアプリケーション別の表示、階層構造の表示をすること、さらにはキーワードの略号を含めることで該略号の音声によるキーワード入力を可能にすること、及びマクロ機能を使ったキーワード表示も含む。

【0021】また、前記音声認識インタフェースは、前記キーワードを制御コマンドに変換するテーブルを有し、このテーブルは1つの制御コマンドに複数のキーワードを対応付け、何れか1つのキーワードの参照で制御コマンドを得るテーブルデータ構成を特徴とする。

【0022】この構成により、利用者がキーワードを正確に覚えていない場合にも、それに近い又は関連するキーワードを発声したときにその音声認識とキーワード対制御コマンドの変換で利用者が意図する制御コマンドが得られるようにする。

【0023】また、前記音声認識インタフェースは、キーワードの分かっている部分の語尾を延ばした発声入力から長音以降を補完して得られるキーワード候補をリスト表示する手段を備えたことを特徴とする。

【0024】この構成により、利用者がキーワードの分かっている部分の語尾を延ばして発声することにより、候補となるキーワードを画面表示し、この画面表示から正しいキーワードを知って、新たに正しいキーワードによる操作指示ができるようにする。

【0025】また、前記音声認識インタフェースは、キーワードの語尾を延ばした長音を伴った音声、又はキーワードに長音を付加した音声を認識したときに、該長音の長さに応じて該キーワードに応じた制御コマンドを繰返し発生する手段を備えたことを特徴とする。

【0026】この構成により、連続的な同じ制御コマンドを繰返し発声するのに代えて、長音を伴う1つのキーワードを発声することにより、同じ制御を繰返し行うのに1回のキーワード発声で済むようにする。

【0027】また、前記音声認識インタフェースは、利用者が音声入力するときのみ音声認識を実行させる音声操作ボタンと、実際の入力音声の認識に適切なレベル以上にあるときのみ音声認識を実行させる入力レベルチェック部と、音声認識結果の信頼度をチェックして正確に認識した音声のみを取り出す認識結果信頼度チェック部とを備えたことを特徴とする。

【0028】この構成により、ノイズの侵入やレベル不足の音声による誤認識を防止し、また認識結果の信頼度チェックもして誤認識を防止する。

【0029】また、本発明は、前記システムソフトウェアは、ポインティングデバイスが指定する位置のウィンドウを利用者が注目する度合いの高いウィンドウとし、当該ウィンドウのフレームレートを高める手段を備えたことを特徴とする。

6

【0030】この構成により、アクティブウィンドウを操作することなく、利用者の興味度の高いウィンドウを容易に指定できるようにする。

【0031】

【発明の実施の形態】

（第1の実施形態）図1は、本発明の一実施形態を示す音声認識インタフェース部のブロック図である。マイクロフォン1は、利用者がコンピュータシステムに操作指示を与えるための音声入力手段にされる。

10 【0032】音声認識インタフェース部2は、音声認識部3と、音声認識用単語辞書4と、キーワード制御コマンド変換部5と、キーワード対制御コマンドテーブル6とを有する。

【0033】音声認識部3は、マイクロフォン1からの音声信号を音声認識用単語辞書4と照合し、認識が一致したキーワードを結果として出力する。例えば、利用者が「開く」と発声した場合、「開く」という文字列をキーワードとして得る。

20 【0034】音声認識用単語辞書4は、図示のように、システムのソフトウェア上で実行されるアプリケーション毎及び各アプリケーションの共通単語が用意され、その中で現在アクティブ状態のアプリケーションの単語と共通単語を有効としている。

【0035】キーワード制御コマンド変換部5は、音声認識部3からのキーワードをキーワード対制御コマンドテーブル6を参照し、該当する制御コマンドを得る。

30 【0036】テーブル6は、各キーワードに1対1に対応づけた制御コマンドが記述され、例えば、キーワード「開く」に対応する制御コマンドがキーボードやマウス等の操作に置き換えられて記述されている。

【0037】コンピュータ本体7のシステムソフトウェア8は、MS-WindowsやX-Window system等のGUI環境を持つソフトウェア構成にされ、キーボード等からの操作指示に応動する機能の他に、キーワード制御コマンド変換部5からの制御コマンドに応動する機能を持つ。

40 【0038】表示部9は、システムソフトウェア8からの表示制御に応じたグラフィックディスプレイ機能を持ち、表示画面9Aには同一画面上で複数のアプリケーション（ウィンドウ）表示を行うための表示制御機能を持つ。

【0039】ここで、本実施形態では、システムソフトウェア8による表示制御手段には、表示画面9Aに表示されるすべてのウィンドウに仮想的なウィンドウ優先順位を付けており、現在アクティブになっているウィンドウを最も優先順位の高いものとしておく。

50 【0040】また、アクティブウィンドウの裏に位置するウィンドウを2番目の優先順位とし、その裏のウィンドウを3番目のウィンドウとするウィンドウ位置に応じた優先順位を設定しておく。

【0041】そして、システムソフトウェア8は、ある特定のキーワード（例えば「ウインドウ」）を設定しておき、その特定のキーワードをキーワード制御コマンド5から受け付けたとき、現在アクティブになっているアプリケーション（ウインドウ）の裏に位置するウインドウ、つまりウインドウ優先順位が2番目のものをアクティブウインドウにするウインドウ切換機能を持つ。

【0042】このウインドウ切換機能によるウインドウ切換時には、それまでアクティブであったウインドウはウインドウ優先順位を最下位にする。

【0043】したがって、本実施形態によれば、目的とするアプリケーションが非アクティブな状態のとき、利用者は音声によってウインドウ切換のキーワードを発声することにより、音声認識部3による音声認識とキーワード制御コマンド変換部5による制御コマンドへの変換がなされ、これによりシステムソフトウェア8による画面9A上のウインドウ切換がなされ、この切換えはウインドウ切換のキーワードが発声される毎に優先順位の高いウインドウに切換えられ、希望するウインドウをアクティブな状態にすることができる。

【0044】これにより、利用者は、操作したいウインドウをアクティブな状態にするためには、マウス等のポインティングデバイス进行操作する事なく、特定のキーワードを発声するのみで済み、これに続けてアクティブな状態のウインドウに対する操作指示を音声で行うことができ、操作効率を高めることができると共に、誤操作も防ぐことができる。

【0045】なお、データの入力にはキーボード等から行う方式、又は音声認識インタフェース部2によるデータの認識と解釈で行う方式のいずれでも良い。

【0046】また、ウインドウの切換えは、特定のキーワード（制御コマンド）を受け付ける毎に優先順位の高い順に切換える場合を示したが、このウインドウ切換方法はシステムソフトウェア側のソフトウェアの変更で適宜変更できる。

【0047】例えば、ウインドウ切換え開始のキーワードを受け付けたときに、ある時間間隔（例えば、1秒毎）でアクティブウインドウを自動的に切換え、ウインドウ切換え停止のキーワードを受け付けたときにウインドウの切換えを停止する方式、さらには本方式と実施形態の方式を組み合わせた方式とすることができる。さらにまた、ウインドウ切換えの指令とウインドウ番号の組を音声で発声し、システム側で当該ウインドウ番号を持つウインドウをアクティブウインドウとする方式でも良い。

【0048】また、グラフィックディスプレイに表示される各ウインドウに優先順位を持たせた「しおり」を付け、この「しおり」を音声で指定することにより任意のウインドウをアクティブにするウインドウ切換方式とすることができる。

【0049】例えば、「しおり」の設定や解除は、ある特定のキーワード、例えば「しおり1の設定」や「しおり1の解除」を利用者が発生することで現在アクティブになっているウインドウに「しおり1」の設定や解除をする。そして、アクティブにするウインドウの指定は、「しおり1」などを発声することにより当該ウインドウをアクティブにする。

【0050】この場合、設定・解除される「しおり」は、アイコン化されて表示され、どの「しおり」を設定・解除するかが一目でわかるようにするのが好ましい。また、「しおり」の使用には音声で行うが、設定や解除はキーボードなど他の入力手段で置換することもできる。

【0051】以上の「しおり」によるウインドウの切換えは、エディタなどのキーワードによる切換に比べて、ウインドウ単位の「しおり」により切換えになって利用者の意図通りのウインドウを1回の発声でアクティブにできる。

【0052】また、システムソフトウェア8にウインドウ管理部を設け、この管理部に表示中のウインドウ名称とアクティブウインドウ名称及びアクティブウインドウ履歴からなるウインドウ情報を管理し、各アプリケーションの最新のアクティブウインドウに切換えることもできる。

【0053】例えば、複数のアプリケーションA～Dが起動され、各アプリケーションの中でも子ウインドウA1、A2が表示されている状態で、音声により「アプリB」と発声すると、管理部がアプリケーションBの最新のアクティブウインドウをサーチして、例えばアプリケーションB1をアクティブにする。

【0054】さらには、音声により「アプリB」を続けて発声すると、次のアプリB2をアクティブにする。

【0055】（第2の実施形態）前記の実施形態において、音声によりウインドウの切換えを行い、当該アプリケーションの操作を行うが、多数のウインドウが重なり合う場合、利用者はどのアプリケーションがアクティブであるかの判断に迷うことが生じる。また、各ウインドウ中にどのような子ウインドウがあるかを利用者が分からなくなる場合がある。

【0056】そこで、本実施形態では、図2に示すように、複数のアプリケーションを起動中に、例えば「タスクリスト」の発声によって、このキーワードに応じた制御コマンドからシステムソフトウェア8は現在起動中のアプリケーションリスト（各ウインドウ名）を図3のようにタスクリストとして表示する構成とする。

【0057】この際、現在アクティブなアプリケーション（例えばアプリケーションA）は、アクティブであることが容易に分かるよう、編み掛け表示や特別な色・枠表示を行う。また、各ウインドウ内にさらに子ウインドウがあった場合、例えばアプリケーションDではその子

ウインドウも階層構造で同時に表示する。また、タスクリストの表示状態でアクティブウインドウを切換え可能にする。

【0058】本実施形態によれば、利用者は現在アクティブなウインドウを覚えておく必要がなく、アクティブウインドウの確認には「タスクリスト」等の特定のキーワードの発声によって複数のウインドウ名を一目で判断でき、容易に認識できる。

【0059】(第3の実施形態)図4は本発明の他の実施形態を示す。同図が図1と異なる部分は、タスク切

換部10を備えた点にある。

【0060】タスク切換部10は、システムソフトウェア8に対して入力デバイスの切換制御を行う。

【0061】通常時は、1つのウインドウ(例えばアプリケーションA)が操作対象となるが、この状態で音声により「タスク切換B」をインタフェース2に与えると、この制御コマンドがタスク切換部10に与えられる。

【0062】このとき、タスク切換部10は、アプリケーションBを音声入力対象とし、アプリケーションAを

【0063】この入力デバイスの切換状態は、利用者に分かるようウインドウ表示画面のうち、アプリケーションAとアプリケーションBは他のアプリケーションと異なり、かつアプリケーションA、B間でも異なる表示、例えば色別表示を行う。

【0064】上記の入力デバイス切換状態からの復帰は、音声により「タスク切換B」を与えることでインタフェース2によるタスク切換部10への復帰コマンドで

【0065】本実施形態によれば、複数のウインドウを操作対象とすることができる。

【0066】(第4の実施形態)図5は本発明の他の実施形態を示す。本実施形態では、キーワード制御コマンド変換部5に特定のキーワードが与えられたときに、単語辞書4に登録されているキーワードリストを一覧表として表示制御できるようにした点にある。

【0067】例えば、図5の構成において、マイクロフォン1を通して、利用者が「キーワード一覧」の音声を発声したとき、音声認識部3ではその音声認識結果をキーワード制御コマンド変換部5に与える。該変換部5は、このキーワードには単語辞書4に登録されているキーワード群を読み出し、それをキーワード一覧表データとすると共にキーワード一覧の制御コマンドをシステムソフトウェア8に与え、システムソフトウェア8による表示制御によりキーワードリストを画面9A上に表示させる。

【0068】図6はキーワードリストの画面表示例を示し、全アプリケーションのキーワード表示と共通単語リ

スト表示なされ、また現在有効なキーワードが分かるよう編み掛け表示を行う。

【0069】また、各アプリケーション内に個別のアプリケーションがある場合は、例えばアプリケーションAのように個別アプリケーション別に単語リストを階層構造表示する。

【0070】なお、キーワードリストの表示を現在有効なアプリケーションのキーワード表示とすることでも良い。

【0071】本実施形態によれば、現在有効なキーワード又は全部のキーワードがリスト形式で表示されるため、利用者はキーワードを覚えておくことなくキーワードリストの呼び出しのためのキーワードを覚えておくことで済む。

【0072】また、階層構造型のキーワードリスト表示になるため、アプリケーション単位のキーワードの認識が容易になる。また、アプリケーション間のキーワードの関連付けが容易になる。

【0073】(第5の実施形態)本実施形態は、前記の実施形態と同様に、キーワード制御コマンド変換部5に特定のキーワードが与えられたときに、現在有効となっているキーワード群を一覧表として表示制御する場合である。

【0074】例えば、マイクロフォン1を通して、利用者が「キーワード一覧」の音声を発声したとき、単語辞書4から現在有効となっているキーワード群を読み出し、それを画面9A上に表示させる。このとき、キーワード群には各々に対応した略号(番号又は記号)が同時に表示される。

【0075】この後の操作指示は、表示されるキーワード群から前述の実施形態と同様にキーワードの音声入力による方法、又はキーワード群の一覧表示に付加する略号(番号又は記号)の音声入力によってなされる。

【0076】図7は、ある監視制御ソフトウェアを用いて「福岡」の監視状況を取り出す場合の操作例を示す。

【0077】まず、利用者は、「キーワード一覧」を発声したときにキーワード制御コマンド変換部5が現在有効なキーワード群を単語辞書4から読み込み、この制御コマンドとキーワード群によってシステムソフトウェア8によりメニュー画面A1として表示する。

【0078】次に、表示されるメニュー画面A1から、利用者は、ファイルのキーワードを知るために「ファイル」を発声する。これにより、ファイルについて有効なキーワード群が単語辞書4から読み出され、ファイル画面A2として表示される。次に、ファイル画面A2から「実行」を選択するには、その「実行」の音声で可能であるが、図示ではその略号としての「1」音声による実行画面A3のキーワード群表示を得る。同様に、実行画面A3から監視制御の略号「2」の音声により監視制御画面A4を得、このうち「福岡」を指定する略号

「4」の音声により「福岡」の監視状況を得ることができる。

【0079】ここで、各画面A1～A4に表示されるキーワード「マクロ」は、音声での発声語と発声順をマクロ化する機能を持たせ、頻度の高いオペレーションには少ないワードの発声で実行できるようにするものである。

【0080】図8は、マクロを使った監視状況を得る場合を示し、メニュー画面A1のマクロの略号「11」の音声入力により、マクロ画面A5を得る。このマクロ画面A5には、地名「福岡」と監視制御の区別「監視」を一度に表示し、このうち略号「3」の音声により「福岡」の監視状況表示を得る。

【0081】以上のことより、本実施形態によれば、キーワードリストを用意することなく、また利用者がキーワードを覚えておくことなく、キーワード一覧表を呼び出すための1つのキーワードを覚えておくことで必要なキーワードを画面表示から容易に知ることができる。

【0082】また、本実施形態によれば、キーワードそのものの以外に、キーワードの略号を受け付けることにより、利用者の発声語数の削減と煩わしさの軽減、これらに伴うシステムの音声認識率の向上を図ることができる。

【0083】また、本実施形態によれば、音声での発声語と発声順をマクロ化することにより、オペレーションの簡略化、高速化、利用者本意の操作指示を実現できる。

【0084】(第6の実施形態)図9は、本発明の他の実施形態を示し、キーワード対制御コマンドテーブルのテーブルデータを示す。

【0085】前述までの実施形態では、特定のキーワードになる音声に反応したウィンドウ切換えや操作指示となるため、このキーワードを利用者が覚えておく必要がある。このため、利用者の記憶が曖昧で正しいキーワードを正確に覚えていないとき、ウィンドウ切換えや操作指示ができない。

【0086】この課題には、キーワードリストを用意することが考えられるが、アプリケーションをGUI環境上で動的に切換えするには各アプリケーション毎のキーワードリストを用意するのは煩わしく、しかも実用的でない。また、アプリケーションに変更が生じた場合にはその都度リストを書き替える必要がある。

【0087】そこで、本実施形態では、キーワードリストを用意する事なく、利用者の曖昧な記憶にも利用者が意図する制御コマンドが得られるように、図示のようなキーワード対制御コマンドテーブル6を用意する。

【0088】図9の例では、ファイルオープンのための制御コマンド記述に対して、複数(図示では3つ)のキーワード「開く」と「開け」及び「オープン」を用意する場合を示す。

【0089】したがって、利用者がファイルオープンを音声で行うには、正しいキーワード「開く」ではなく、「開け」や「オープン」を連想してそれを発声した場合、キーワード制御コマンドテーブル6を参照した変換部5によって正しい制御コマンドへの変換がなされ、目的とする制御がなされる。同様に、ウィンドウ切換えも複数のキーワードを用意しておくことでなされる。

【0090】本実施形態によれば、利用者が正しいキーワードを覚えておく必要がなく、またキーワードリストを用意する事なく、操作に関連して連想できる近い言葉を発声することで所期のキーワードによる操作指示を得ることができる。

【0091】なお、テーブル6への複数のキーワードの追加や変更などの編集には、キーボードからの入力でも可能であるが、音声による編集でも良い。

【0092】この音声による編集は、例えば、システムソフトウェア8への制御コマンドとして「テーブル6の編集」を音声で与えることで該テーブルを編集可能な状態にし、次にキーワード編集のキーワード(例えば、「登録」の音声)を発声し、正しいキーワードの発声によりテーブル6から該キーワードを持つ部分を検索し、次いで別名のキーワードの発声により該別名のキーワードをテーブルに登録することで実現される。

【0093】(第7の実施形態)これまでの実施形態では、少なくとも1つのキーワードを正確に覚えておく必要があるが、利用者の記憶が曖昧なため、正しいキーワードを忘れてしまい、キーワードの途中まで分かっている場合がある。

【0094】そこで、本実施形態では、キーワードの途中まで分かっている部分から目的とするキーワードを得るシステムを提供する。

【0095】本実施形態では、利用者が途中まで分かっている部分のキーワードの音声入力を行い、この音声入力に分かっている部分の語尾を延ばして発声することにより、長音以降を補完して得られる正しいキーワード候補のリストを表示する構成とする。

【0096】例えば、ファイルの内容を表示するためのキーワードを「表示」とした場合、テーブル6に登録されているキーワード「表示[ひょうじ]」に対して、「ひょう」まで分かっているならば「ひょうー」と語尾を延ばした発声を行う。

【0097】この語尾が長音のキーワードを受け取ったキーワード制御コマンド変換部5は、テーブル6内のキーワードを検索し、その認識キーワード「ひょう」で始まる正しいキーワード候補のリスト(例えば、表示、標準、評価、…)の画面表示コマンドをシステムソフトウェア8に与え、その画面表示を得る。

【0098】この画面表示から、利用者が正しいキーワード「表示」を改めて発声することにより、正しいキーワード「表示」による操作指示を得る。

【0099】本実施形態によれば、利用者が正確なキーワードの語尾を忘れている場合にもその初めの部分を見ていれば、システムが表示したキーワードリストの中から正しいキーワードを見つけだし、正しいキーワードの音声入力ができる。

【0100】(第8の実施形態)これまでの実施形態から、正しいキーワードの音声入力を知って音声入力を行うにおいて、キーワードとして画面の「拡大」、「縮小」、上下左右への「移動」のような連続的な操作指示を行う場合、利用者は期待する画面状態になるよう、キーワードを繰り返し発声する(例えば、「拡大」、「拡大」、「拡大」という煩わしさがある。

【0101】そこで、本実施形態では、連続した同じ操作指示のキーワード入力を簡易にするシステムを提供する。

【0102】本実施形態では、キーワードの語尾を延ばした長音を伴った発声をしたときに、この長音の長さに応じて連続的な同じ制御コマンドを繰り返し得ることにより、同じキーワードの繰り返し発声を不要にする。

【0103】例えば、画面の拡大を行うキーワード「拡大」の発声に、長音を伴った「かくだーい」と発声する。この発声に対して、音声認識部3は、音声「かくだ」の部分から、単語辞書4の参照で拡大動作であることを認識し、キーワードとして「かくだい」とこれが連続することの符号をキーワード制御コマンド変換部5に与え、該変換部5によって画面拡大の制御コマンドの発生を繰り返す。

【0104】そして、音声認識部3がキーワード「拡大」の語尾になる「い」を認識したときに繰り返し制御の停止を意味する符号を変換部5に与え、これにより変換部5が「拡大」の制御コマンドの発生を停止し、システムソフトウェア8が画面の拡大表示制御を停止する。

【0105】画面の縮小や上下左右への移動についても同様になされ、さらにプルダウンメニュー内におけるメニュー選択などにも利用することができる。すなわち、プルダウンメニュー内の項目を選択するとき、カーソルの上下移動のそれぞれのキーワード「うえ」、「した」に長音を伴って発生することにより、長音が続く限りカーソルを上下方向に移動させ、発声が終了したときにカーソル移動が停止し、停止位置の項目選択を可能にする。

【0106】本実施形態によれば、利用者は、同じキーワードによる連続した操作指示を行うには、キーワードとして長音を伴った発声を行うことにより、長音が続く間だけ画面の拡大等の連続制御がなされ、希望する画面状態になったときに発声を終了することにより連続制御が停止され、利用者が期待する画面状態を得ることができる。これにより、利用者はキーワードの発声回数を1回で済み連続した操作指示が容易になる。

【0107】なお、実施形態において、長音を伴ったキ

ーワードに代えて、語尾に長音を付加したキーワード(例えば「かくだーい」)を発声し、この長音が続く限り当該キーワードの操作指示を繰り返す構成にして同等の作用効果を奏する。

【0108】(第9の実施形態)図1や図4、5のシステム構成において、一般に音声認識は100%の認識率が達成されないため、システムの誤操作の一因となる場合が予想される。

【0109】そこで、本実施形態では、音声認識に誤りが発生したときにも誤操作を確実に防止するシステムを提供する。

【0110】図10は、本実施形態の音声認識インタフェースを示す。音声操作ボタン11は、ソフトウェアボタンにされ、このオン操作で音声入力の許可信号を音声入力インタフェース12に与える。このボタン11を設けることにより、ボタン11のオフ時には音声入力を遮断し、周囲の騒音等によるノイズが音声認識部3に侵入するのを無くしてノイズによる誤操作を防止する。

【0111】入力レベルチェック部13は、インタフェース12を通した音声信号のレベルをチェックし、入力レベルが通常より低い音声信号を遮断することにより、低いレベルの音声認識による誤り発声を防止する。

【0112】また、入力レベルチェック部13は、音声信号を遮断したときに、応答機能部14に遮断したことの信号を渡す。

【0113】認識結果信頼度チェック部15は、音声認識部3による認識結果の信頼度をチェックし、信頼度が低いときには認識結果を変換部5に渡すのをやめる。これにより、音声認識をその信頼度チェックにより誤った音声による操作を防止する。

【0114】また、チェック部15は、信頼度が低いときに応答機能部14に操作を中止したことの信号を渡す。

【0115】応答機能部14は、チェック部13又は15からの音声認識異常の信号を受け取ったときに、利用者への再発声を要求する。また、この要求に対する再発声にも異常の信号を受けたときに音声操作ボタン11を強制オフすると共に、利用者ひ他の入力インタフェース(マウスやキーボードなど)による入力要求を出す。

【0116】したがって、本実施形態によれば、音声操作ボタン11により利用者が音声入力するときのみ音声認識を実行させ、入力レベルチェック部13により実際の入力音声認識が認識に適当なレベル以上にあるときのみ音声認識を実行させ、認識結果信頼度チェック部15により音声認識結果の信頼度をチェックして正確に認識した音声のみに応答できるようにする。

【0117】また、音声入力及び認識の異常発生には、利用者に再入力を促したり、他のインタフェースによる入力を促すことにより、確実な音声入力と認識を得ることができる。

15

【0118】(第10の実施形態)複数のウィンドウを同時に表示するにおいて、利用者が注目しているウィンドウの度合い(興味度)の高さに応じて各ウィンドウのフレームレートを変化させ、限られたシステムリソースの最適配分を図ることができる。

【0119】この興味度の高さは、ウィンドウの表示状態から推測することができる。そのルールとしては、図11の(a)に示すように、アクティブウィンドウに近いウィンドウほど興味度が高いとし、またウィンドウの面積が小さいほど興味度が低いとすることができる。興味度の度合いによりウィンドウのフレームレートを変えることも行われる。

【0120】このルールでは、ウィンドウの表示状態により興味度が自動的に決定され、利用者が任意に興味度を変えるためにアクティブウィンドウを変えるなどの操作を必要とし、操作が複雑になる。

【0121】そこで、本実施形態では、利用者の興味度の度合いを確実にかつ自動的に推測して最適配分の表示を得るシステムを提供する。

【0122】本実施形態では、図11の(b)に示すように、マウス等のポインティングデバイスが位置するウィンドウを利用者の興味度が高いウィンドウとシステムソフトウェアが認識し、またそのウィンドウのフレームレートを高める。このとき、アクティブウィンドウは変更しない。

【0123】したがって、本実施形態では、マウス等のポインティングデバイスのポインタ位置を変えることにより、自動的にウィンドウの興味度を変え、またフレームレートも変えるため、アクティブウィンドウの変更操作をすることなく、利用者が任意に特定のウィンドウの興味度を上げることができる。このため、利用者の興味度のあるウィンドウのフレームレートを上げる操作が単純化され、より使い易いシステムを構築できる。

【0124】

【発明の効果】以上のとおり、本発明によれば、音声認識インタフェースを利用し、音声によるウィンドウ切換えを行うようにしたため、マウス等による従来のウィンドウ切換え操作を不要にし、音声によるウィンドウ切換えになって操作効率を高めると共に、誤操作を防止することができる。

【0125】また、本発明は、複数のアプリケーションのウィンドウ情報を管理し、音声により各アプリケーションを指定したときにウィンドウ情報からアプリケーションの最新のアクティブウィンドウに切換え、同じアプリケーション名の音声により当該アプリケーションの新しいウィンドウ順に切換えるようにしたため、アプリケーションのウィンドウのうち良く利用するウィンドウ指定を容易にする効果がある。

【0126】また、本発明は、音声入力により、アクティブなウィンドウを他のウィンドウから区別した表示を

16

行うこと及びアクティブなアプリケーションを区別したタスクリスト表示を行うようにしたため、利用者がアクティブなウィンドウ及びアプリケーションを認識するのを容易にする効果がある。

【0127】また、本発明は、タスク切換により、音声によるアプリケーションへの入力と、マウス等の他の入力デバイスによる他のアプリケーションへの入力とするため、複数のウィンドウを操作対象とすることができる。

10 【0128】また、本発明は、利用者がキーワード一覧等の特定のキーワードを発声したときに単語辞書に登録されるキーワード又は現在有効なキーワードリストを画面表示するため、この画面表示から利用者が必要とするキーワードを容易に知ることができ、利用者はキーワード一覧等の1つのキーワードを覚えておくのみで有効なキーワードの全てを画面表示から知ることができる。

【0129】また、本発明は、キーワードを制御コマンドに変換するテーブルには1つの制御コマンドに複数のキーワードを対応付け、何れか1つのキーワードから制御コマンドを得るテーブルデータ構成とするため、利用者がキーワードを正確に覚えていない場合にも、それに近い又は関連するキーワードの発声により利用者が意図する制御コマンドが得られ、キーワードの正確な記憶を不要にして使い勝手を良くする効果がある。

【0130】また、本発明は、利用者がキーワードの分かっている部分の語尾を延ばして発声することにより、候補となるキーワードを画面表示するようにしたため、利用者はキーワード候補の画面表示から正しいキーワードを知ることができ、キーワードの一部を覚えておけば  
30 正しいキーワードを知ることができる。

【0131】また、本発明は、連続的な同じ制御コマンドを繰り返し発声するのに代えて、長音を伴う1つのキーワードを発声することにより、長音が続く限り同じ制御を繰り返し行うようにしたため、同じキーワードによる操作指示には当該キーワードに長音を付加した1回の発声で済み、連続した同じ操作指示が簡単になる。

【0132】また、本発明は、利用者が音声入力するときのみ音声認識を実行させ、実際の入力音声認識に適当なレベル以上にあるときのみ音声認識を実行させ、音声認識結果の信頼度をチェックして正確に認識した音声のみに応答できるようにするため、音声認識の誤りによる誤操作を確実に防止できる。また、音声入力及び認識の異常発生には、利用者に再入力を促したり、他のインタフェースによる入力を促すことにより、確実な音声入力と認識を得ることができる。

【0133】また、本発明は、マウス等のポインティングデバイスが位置するウィンドウを利用者の興味度が高いウィンドウとシステムソフトウェアが認識し、ウィンドウのフレームレートを高めるようにしたため、アクティブウィンドウの変更操作をすることなく、利用者が任  
50

意に特定のウィンドウの興味度を上げることができ、利用者の興味度のあるウィンドウのフレームレートを上げる操作が単純化され、より使い易いシステムを構築できる効果がある。

【図面の簡単な説明】

【図1】本発明の一実施形態を示すブロック図。

【図2】他の実施形態におけるウィンドウイメージ。

【図3】他の実施形態におけるタスクリストウィンドウイメージ。

【図4】他の実施形態のブロック図。

【図5】他の実施形態のブロック図。

【図6】他の実施形態におけるキーワードリストウィンドウイメージ。

【図7】他の実施形態における音声による操作指示手順図。

【図8】他の実施形態における音声による操作指示手順図。

【図9】他の実施形態におけるキーワード対制御コマンドテーブルのデータ構成例。

【図10】他の実施形態における音声認識インタフェース

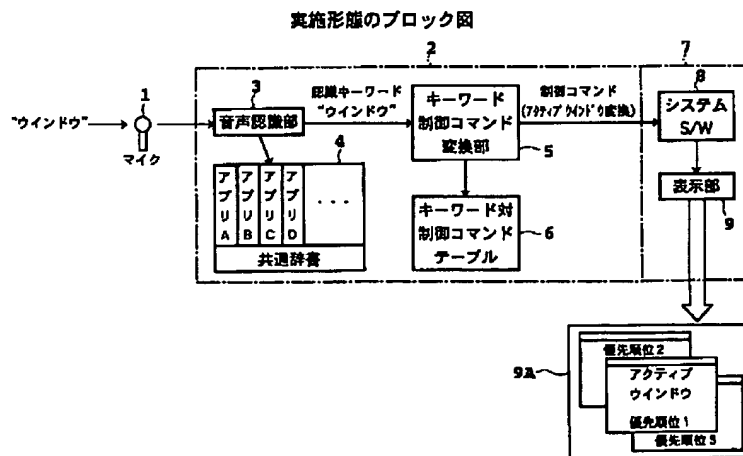
の構成例。

【図11】他の実施形態における興味度によるウィンドウ表示例。

【符号の説明】

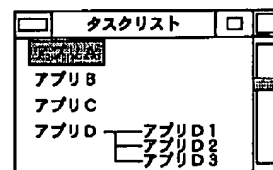
- 1…マイクロホン
- 2…音声認識インタフェース
- 3…音声認識部
- 4…単語辞書
- 5…キーワード制御コマンド変換部
- 6…キーワード対制御コマンド変換テーブル
- 7…コンピュータ本体
- 8…システムソフトウェア
- 9…表示部
- 10…タスク切替部
- 11…音声操作ボタン
- 12音声入力インタフェース
- 13…入力レベルチェック部
- 14…応答機能部
- 15…認識結果信頼度チェック部

【図1】



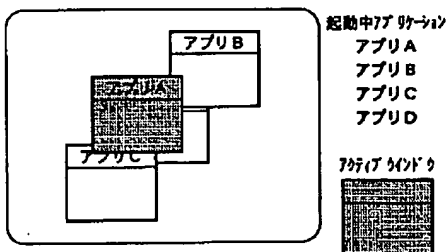
【図3】

タスクリストウィンドウイメージ



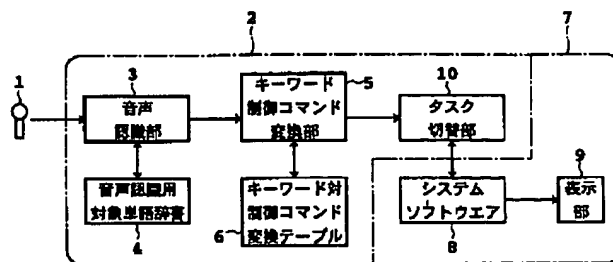
【図2】

ウィンドウイメージ



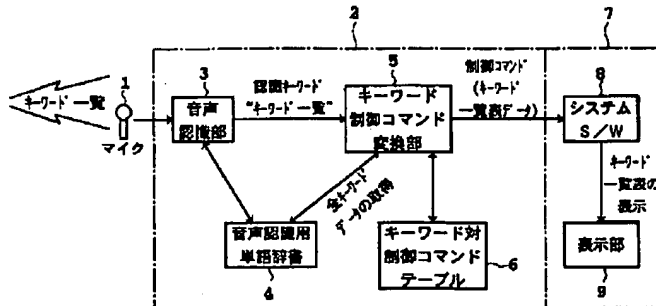
【図4】

他の実施例のブロック図



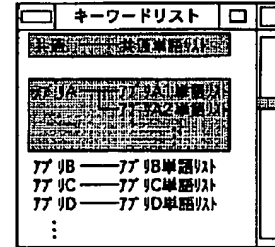
【図5】

他の実施例のブロック図



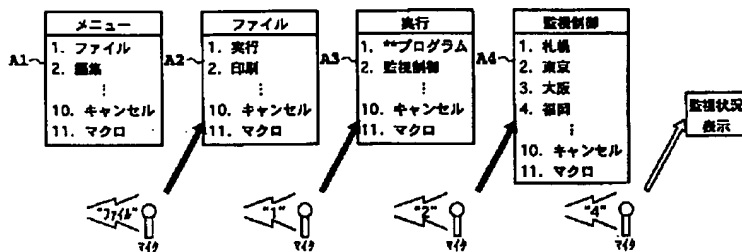
【図6】

キーワードリストウィンドウイメージ



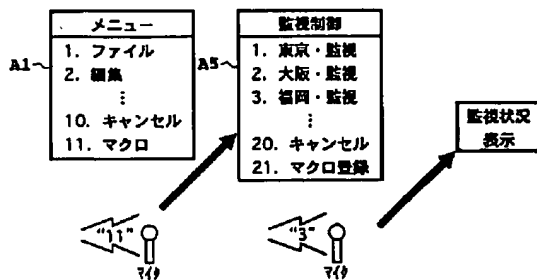
【図7】

音声による操作指示手順図



【図8】

音声による操作指示手順図



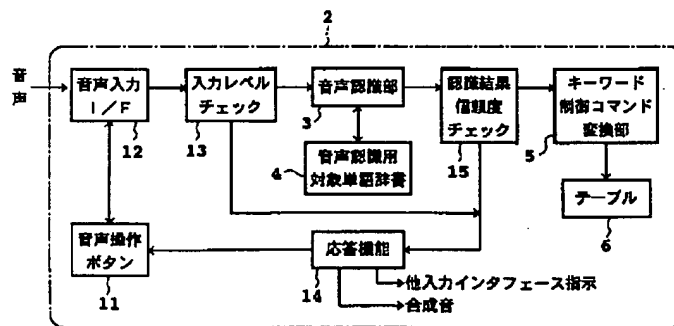
【図9】

テーブルデータ構成例

キーワード1	キーワード2	キーワード3	制御コマンド記述
開く	開け	オープン	ファイルオープンのためのコマンドの記述
...	...	...	.....
...	...	...	.....
...	...	...	.....
...	...	...	.....

【図10】

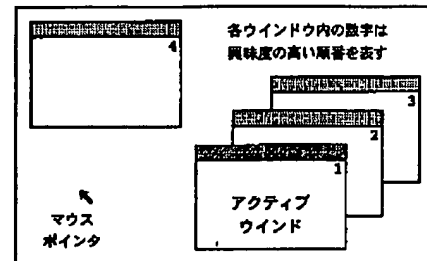
音声認識インタフェースシステム構成例



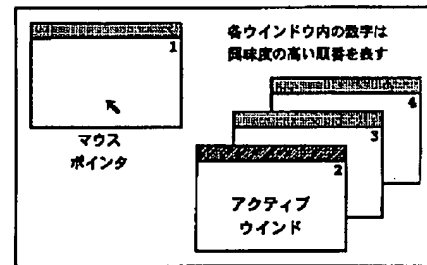
【図11】

興味度の決定例

(a)



(b)



フロントページの続き

(72)発明者 原 洋  
東京都品川区大崎2丁目1番17号 株式会  
社明電舎内

(72)発明者 松田 啓寿  
東京都品川区大崎2丁目1番17号 株式会  
社明電舎内

(72)発明者 小形 毅  
東京都品川区大崎2丁目1番17号 株式会  
社明電舎内

(72)発明者 望月 燕雅  
東京都品川区大崎2丁目1番17号 株式会  
社明電舎内

(72)発明者 池田 勇  
東京都品川区大崎2丁目1番17号 株式会  
社明電舎内

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**